



Novel scheme for efficient and cost-effective forwarding of packets in optical networks without header modification

Wessing, Henrik; Fjelde, Tina; Christiansen, Henrik Lehrmann; Dittmann, Lars

Published in:

Proceedings on Optical Fiber Communication Conference and Exhibit

Link to article, DOI:

[10.1109/OFC.2001.927511](https://doi.org/10.1109/OFC.2001.927511)

Publication date:

2001

Document Version

Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Wessing, H., Fjelde, T., Christiansen, H. L., & Dittmann, L. (2001). Novel scheme for efficient and cost-effective forwarding of packets in optical networks without header modification. In *Proceedings on Optical Fiber Communication Conference and Exhibit* (Vol. 4) <https://doi.org/10.1109/OFC.2001.927511>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Novel scheme for efficient and cost-effective forwarding of packets in optical networks without header modification

H. Wessing, T. Fjelde, H. Christiansen and L. Dittmann

Research Center COM, Techn. Univ. of Denmark, Bldg 349, DK-2800 Lyngby, Denmark
Telephone: +45 4525 3804, Fax: +45 4593 6581, E-mail: hw@com.dtu.dk

Abstract: We present a novel scheme for addressing the outputs in optical packet switches and demonstrate its good scalability. The scheme requires neither header modification nor distribution of routing information to the packet switches, thus reducing optical component count while simplifying network management.

© 2001 Optical Society of America

OCIS codes: (060.4250) Networks; (999.9999) Packet forwarding

1 Introduction

With the exponential growth in Internet traffic, scaling of IP data networks far beyond their present performances is required. In order to meet this challenge, Multi-Protocol Label Switching (MPLS) [?] is rapidly emerging as the technology that allows for enhanced speed and scalability as well as service provisioning capabilities in the Internet. All-optical packet switching fits well into the MPLS concept, while offering flexibility, enabling scaling to very high bit rates and eliminating potential electronic bottlenecks.

Packet forwarding within MPLS is traditionally based on label swapping, wherein the label in an incoming packet is replaced with a new value in the packet switch as it is forwarded to the appropriate outlet. Thus, exploiting the advantages of optical MPLS would traditionally require all-optical header modification (i.e., label swapping) in the packet switches, as demonstrated in, e.g., [?][?]. However, as this added functionality complicates the packet switch implementation, while increasing the optical component count, it would be advantageous to avoid.

In this paper, we present a scheme, which conforms with the MPLS concept, and which allows the decoding of a fixed label within each packet switch, thereby enabling the extraction of the desired forwarding information without changing the header. The scheme is based on an arithmetic operation in the core nodes, which can easily be performed within the duration of the packet. The good scalability of the scheme, which is important for practical implementation, is demonstrated through simulations for different network sizes and node size distributions.

2 Key identification scheme

The principle of packet forwarding within a network can basically be divided into three groups based on the operating principle. The first is an approach where a path through the network is preestablished and designated by a globally unique identification number, an approach much like IP routing. However, this method requires very large lookup tables in the core nodes and global distribution of forwarding information, which is time consuming. The second approach is based on label swapping, where a label is replaced or modified for each node using an entry in a lookup table. This method reduces the lookup tables but time consuming label distribution is still required, which increases the management complexity. Furthermore, optical components are required to perform the header modification, which is undesirable. The third approach is to carry all the required packet forwarding information within the header. Using this approach avoids the demand for label distribution as the forwarding information is carried solely in the header. However, whether or not header modification can be avoided in this case depends on the scheme that is used. If a label stack scheme is employed, where labels are stored as a stack and the top label is popped or the label stack is rotated at each node, then some sort of header modification is required. Header modification can be avoided, if it is possible to encode the forwarding information in the label such that each passed node can decode it differently and extract the desired output port. The key identification scheme (KIS), that we propose in this paper, satisfies these requirements. The scheme uses modular arithmetic operations to decode a fixed header and extract a desired number (here the output port) depending on a node specific identification number, which is denoted the *key*. The mathematical algorithm we use is the old Chinese Remainder Theorem (CRT), which states that an integer scalar a can be computed with the following property, which is derived in [?]:

$$a_i = a \bmod n_i \quad \text{for } i = \{1, \dots, k\}$$

where a_i and n_i are entries in the arrays \bar{a} and \bar{n} , respectively, and k is the number of traversed nodes. Each traversed node uses a pair of integers (n_i, a_i) that is defined as follows: n_i uniquely identifies the node in the network and a_i is the desired output at that node. The scalar a , which we denote the Route Designation Field (RDF), is generated prior to launching the packet into the network, and hereafter the packet is forwarded at each node by decoding a_i using the key n_i and equation (1). The keys (n_i) must be pairwise relative primes, i.e., the greatest common divisor (gcd) for any pair of keys in the network must be one. Furthermore, the key should be larger than the desired output, i.e., $n_i > a_i$ for all i .

In figure 1 is given an example that shows a packet traversing three nodes that are identified by the unique keys, 25, 14 and 37. Note that the keys are pairwise relative primes even though only 37 is an absolute prime. Considering the figure, it is possible to construct the two arrays $\bar{n} = (25, 14, 37)$ and $\bar{a} = (4, 2, 3)$, and the scalar a can be computed using an algorithm given in [?] which gives $a = 2704$. This scalar is carried in the header and at the three nodes the following calculations are performed: $a_1 = 2704 \bmod 25 = 4$, $a_2 = 2704 \bmod 14 = 2$ and $a_3 = 2704 \bmod 37 = 3$. Note that it is possible to cascade as many nodes as desired, by allowing for a sufficiently large RDF.

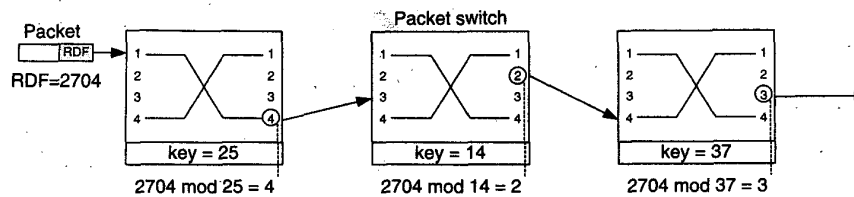


Fig. 1. The key identification scheme used to switch a packet to output 4, 2 and 3 at nodes with keys 25, 14 and 37, respectively

The implementation of the key identification scheme is quite simple as each node is applied a fixed key, which is only changed in response to larger reconfigurations of the network, e.g., an increase in the network size. Hereafter, every path through the network can be chosen at the network edge nodes by computing the RDF using the Chinese Remainder Theorem. In case of congestion, fiber break, etc., the path is simply changed by recomputing the RDF. Thus, it is not necessary to use label distribution protocols to distribute labels. The operation performed in the electrical control system of the optical node is thus a simple modular expression, which generally can be calculated by a division and a subtraction. Furthermore, as the key is fixed, the inverse key can be calculated initially and the division is thus replaced with a multiplication, that can be carried out with very limited time consumption of only a fraction of the packet duration time even if the packet length is only, e.g., about $1 \mu\text{sec}$. Hence, the key identification scheme has significant advantages as implementation is easily performed electrically, a control system to update forwarding tables throughout the network is not needed, and costly optical devices to perform label swapping are avoided.

3 Scalability of the scheme

Another important factor for implementing the scheme is the scalability when the network size increases. In the following the key identification scheme is compared with the label stack scheme as they both do not require label distribution (group three mentioned above). The scalability of the key identification scheme relies on the available number of pairwise relative primes in the network. If the network has a number of nodes with different sizes then each node is just applied a key that is larger than or equal to its size, i.e., it is not necessary to dimension all the keys in the network for the largest node size. The label size for a given path will, according to [?], always be smaller than the product of the keys for the traversed nodes. This means that with a given label size, the key identification scheme can as well support long routes and small nodes as short routes and larger nodes. On the contrary, the label stacking scheme must be predesigned for the worst case, that is, the largest node size in the network and the longest allowed route. Thus, it can be expected that the KIS improves its performance compared to the label stack scheme when operating on a mixed network. This is demonstrated in figure 2 where randomly generated networks are evaluated by finding the shortest route between every pair of nodes using Dijkstra's algorithm. Figure 2(a) is computed for a network with a rather uniform distribution of node sizes with 50% 8×8 nodes and 50% 16×16 nodes, while figure 2(b) is computed for a much more mixed network where the node size distribution is as follows: 10% 4×4 , 20% 8×8 , 30% 16×16 , 20% 32×32 , 10% 64×64 , 5% 128×128 and 5% 256×256 nodes.

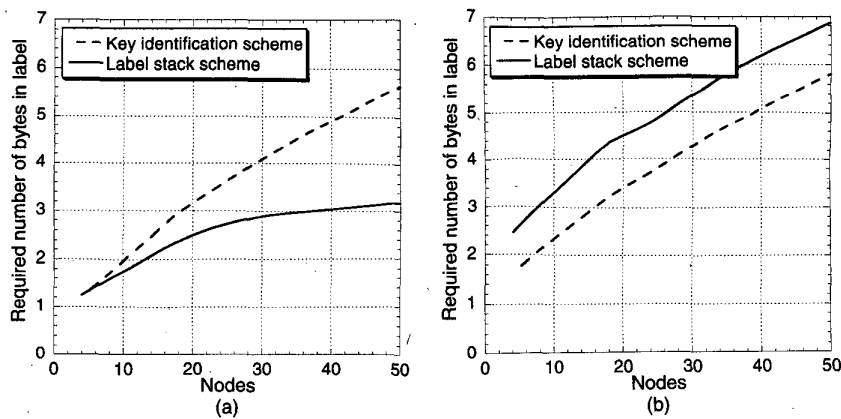


Fig. 2. Scalability of the two packet forwarding schemes. (a) is for a network with only two node sizes and (b) is for a network with up to seven different node sizes.

By comparing the figures it is seen that the key identification scheme requires a label size, of about six bytes for network sizes up to about 50 nodes in the uniform case, which is approximately three bytes more than for the label stack scheme. However, for a mixed network the situation is different and the KIS requires the same label size, whereas the label stack scheme requires up to seven bytes in the label field for 50 nodes in the network. In both cases route lengths of nine cascaded nodes are employed as the maximum length.

It is seen that the key identification scheme requires approximately the same number of bytes for a very mixed network as for a uniform network. This demonstrates a superior flexibility compared with the label stack scheme, however, the scalability of a uniform network is inferior. This, however, is most likely a minor problem for actual implementation as one would expect a mixed network in practice.

It is obvious that both the KIS and the label stack scheme requires a larger label field than the label swapping scheme, which can be as small as 4 bytes [?]. However, we believe that the advantages associated with the key identification scheme far outweigh the disadvantage of the limited reduction in the bandwidth efficiency.

4 Conclusion

In this paper, a novel scheme for packet forwarding in an optical packet switched network is presented. The scheme relies on the encoding of forwarding information in the packet label such that each node can decode it differently and extract the desired output port. To perform this we use an algorithm based on the Chinese Remainder Theorem. The scheme has significant advantages compared to the well-known label swapping, as it does not require time-consuming distribution of packet forwarding information throughout the network, and avoids the use of optical components for header modification. Thus, the component count in packet switches is reduced, increasing the cost-effectiveness.

The scalability of the scheme is evaluated, demonstrating that the size of the required label field scales well with the network size and is independent of node size variations, thus indicating the flexibility of the scheme. For networks consisting of up to 50 nodes, we found that less than seven bytes are required in the label field. As approximately four bytes are required for traditional label swapping, the scheme has a somewhat lower bandwidth efficiency. However, due to the significant advantages that are otherwise associated with the scheme, we believe that it has a great potential for use in future optical networks.